Single-molecule insight into target recognition by CRISPR-Cas complexes

Authors

Marius Rutkauskas[1], Andrey Krivoy[1,2], (Konstantin Severinov), Mark Szczelkun[3], Christophe Rouillon[1,4]*, Ralf Seidel[1]*

[1] Molecular Biophysics group, Institute for Experimental Physics I, Universität Leipzig, Leipzig, Germany

[2] Skolkovo Institute of Science and Technology, Skolkovo, Russia

[3] DNA-Protein Interactions Unit, School of Biochemistry, University of Bristol, Bristol BS8 1TD, UK

[4] Biomedical Sciences Research Complex, University of St Andrews, Fife KY16 9ST, UK

* Corresponding author

ABSTRACT

Ribonucleoprotein (RNP) complexes from CRISPR-Cas systems have attracted enormous interest since they can be easily and flexibly reprogrammed to target any desired locus for genome engineering and gene regulation applications. Basis for the programmability is a short RNA (crRNA) inside these complexes that recognizes the target nucleic acid by base pairing. For CRISPR-Cas systems that target double-stranded DNA this results in local DNA unwinding and formation of a so-called R-loop structure. Here we provide an overview how this target recognition mechanism can be dissected in great detail at the level of a single molecule. Specifically, we demonstrate how magnetic tweezers are applied to measure the local DNA unwinding at the target in real time. To this end we introduce the technique and the measurement principle. By studying modifications of the consensus target sequence we show how different sequence elements contribute to the target recognition mechanism. From these data a unified target recognition mechanism can be concluded for the RNPs Casacade and Cas9 from Type I and Type II CRISPR-Cas systems. R-loop formation is hereby initiated on the target at an upstream element, called protospacer adjacent motif (PAM), from which the R-loop structure zips directionally towards the PAM-distal end of the target. At mismatch positions the R-loop propagations stalls and further propagation competes with collapse of the structure. Upon full R-loop zipping conformational changes within the RNPs trigger degradation of the DNA target. This represents a shared labour mechanism in which zipping between nucleic acid strands is the actual

target recognition mechanism while sensing of the R-loop arrival at the PAM distal end just verifies the success of the full zipping.

I. Introduction

CRISPR-Cas systems have only been discovered within the past decade where they attracted considerable interest (van der Oost, Westra, Jackson & Wiedenheft, 2014). The molecular understanding of some of their enzymatic components, e.g. the Cas9 protein, has been exploited to develop new tools for genome engineering and gene regulation, that are more easy to generate than existing technologies such as ZFNs and TALENs (Gilles & Averof, 2014). CRISPR-Cas systems are present in most archaea and in 10-40% of the bacteria (Burstein, Sun, Brown, Sharon, Anantharaman, Probst, et al., 2016). They are typically referred as adaptive and heritable immune systems in the sense that microorganisms acquire resistance to extrachromosomic elements, such as viruses or plasmids (Bolotin, Quinquis, Sorokin & Ehrlich, 2005, Mojica, Diez-Villasenor, Garcia-Martinez & Soria, 2005, Pourcel, Salvignol & Vergnaud, 2005). This is achieved by integrating short DNA sequences (<40 bp) into the CRISPR loci in their genomes that act as memory of former infections (Barrangou, Fremaux, Deveau, Richards, Boyaval, Moineau, et al., 2007, Mojica, Diez-Villasenor, Garcia-Martinez & Almendros, 2009) (Fig. 1). This step still poorly understood at the molecular level is called adaptation. Each integrated sequence (also called spacer)(Grissa, Vergnaud & Pourcel, 2007) is separated from the next spacer by a short identical repeat that often are palindromic (Lawrence & White, 2011). The name of CRISPR was first derived from the discovery of these islands, i.e. clusters of regularly interspaced repeats (Ishino, Shinagawa, Makino, Amemura & Nakata, 1987). Following transcription of the CRISPR locus, the repeats are specifically recognised by a ribonuclease to generate small RNAs, also called crRNAs (for CRISPR-RNAs) (Charpentier, Richter, van der Oost & White, 2015). The crRNAs are then integrated into large monomeric or multimeric protein complexes formed by the CRISPR associated proteins (Cas proteins), which scan the cellular nucleic acids for the presence of a target sequence. When a nucleic sequence complementary to the crRNA is encountered, it will be degraded either by the ribonucleoprotein (RNP) complex itself or by recruitment of an additional factor displaying the nuclease activity. This stage is called interference and the target sequence is named the protospacer, referring to a previously encountered DNA sequence.

Bioinformatics analysis of the Cas proteins allowed the classification of the CRISPR systems in different types and subtypes. The last published classification proposed up to 5 types (from type I to V) (Makarova, Wolf, Alkhnbashi, Costa, Shah, Saunders, et al., 2015) among which the type I, II and III represent the best studied systems (van der Oost et al., 2014). Common feature for all types is the presence of the proteins Cas1 and Cas2, which are involved in the capture and integration of new

spacers in the adaption stage (Makarova, Haft, Barrangou, Brouns, Charpentier, Horvath, et al., 2011, Yosef, Goren & Qimron, 2012) as well as the presence of a crRNA containing RNP complex for target recognition at the interference stage. By integrating crRNAs with a specifically designed sequence these RNP complexes can be reprogrammed to recognize practically any target of choice. The different types of CRISPR-Cas systems use different RNP complexes and further distinguish themselves by the presence of a specific "signature protein" that is responsible for DNA degradation which is respectively Cas3, Cas9 and Cas10 for the Type I, II and III (see Figure 1). Type I systems employ a large multi-subunit RNP complex called Cascade that recognizes double-stranded DNA (dsDNA) targets. After target recognition and verification, Cascade recruits the signature protein Cas3 - a fused helicase-nuclease – to degrade DNA (Sinkunas, Gasiunas, Fremaux, Barrangou, Horvath & Siksnys, 2011, Sinkunas, Gasiunas, Waghmare, Dickman, Barrangou, Horvath, et al., 2013). In type II systems, the monomeric Cas9 protein is both the RNP for dsDNA target recognition as well as the signature nuclease for target degradation. Using its two nuclease domains it readily generates a double strand break on bound targets (Gasiunas, Barrangou, Horvath & Siksnys, 2012, Jinek, Chylinski, Fonfara, Hauer, Doudna & Charpentier, 2012). It represents a minimal system and became therefore the preferred tool in CRISPR-Cas-based genome engineering applications (Hsu, Lander & Zhang, 2014, Karvelis, Gasiunas, Miksys, Barrangou, Horvath & Siksnys, 2013, Mali, Yang, Esvelt, Aach, Guell, DiCarlo, et al., 2013, Sander & Joung, 2014). In type III systems the RNP complex is multimeric with a similar helicoid structure as found for Cascade (Cas7 family proteins) (Benda, Ebert, Scheltema, Schiller, Baumgartner, Bonneau, et al., 2014, Rouillon, Zhou, Zhang, Politis, Beilsten-Edmands, Cannone, et al., 2013). Despite this similarity the RNP complex is not recognizing dsDNA but complementary RNA sequences (Hale, Zhao, Olson, Duff, Graveley, Wells, et al., 2009, Tamulaitis, Kazlauskiene, Manakova, Venclovas, Nwokeoji, Dickman, et al., 2014, Zhang, Rouillon, Kerou, Reeks, Brugger, Graham, et al., 2012). RNA recognition stimulates a non-specific DNA cleavage activity of the Cas10 signature nuclease that is part of the RNP complex (Elmore, Sheppard, Ramia, Deighan, Li, Terns, et al., 2016, Estrella, Kuo & Bailey, 2016, Kazlauskiene, Tamulaitis, Kostiuk, Venclovas & Siksnys, 2016), such that DNA cleavage is achieved cotranscriptionally (Kazlauskiene et al., 2016, Samai, Pyenson, Jiang, Goldberg, Hatoum-Aslan & Marraffini, 2015).

The central and most crucial step during interference and genome editing is the recognition and the verification of the target sequence by the RNP complex. The recognition should be specific enough to avoid degradation of undesired targets (off-targets). All wild-type CRISPR-Cas systems found so far are, however, somewhat promiscuous, i.e. they tolerate a number of mismatches between crRNA and target (Fineran, Gerritzen, Suarez-Diez, Kunne, Boekhorst, van Hijum, et al., 2014). This is suggested to be beneficial during the defence of foreign DNA, since invaders (such as viruses) can less easily escape by mutations in the protospacer. Promiscuous target recognition leading to

massive off-targeting is however highly problematic in genome engineering applications (Wang, Wang, Wu, Wang, Qiu, Chang, et al., 2015, Wu, Scott, Kriz, Chiu, Hsu, Dadon, et al., 2014). It is therefore crucial to understand the mechanism of target recognition and to develop experiments with which quantitative insight into this process can be obtained.

Most of the knowledge how CRISPR-Cas systems recognize protospacer targets has been gained so far for Type I and II systems. They both recognize dsDNA. In addition to a well matching target that is complementary to the crRNA they require a short nucleotide motif upstream of the protospacer, called PAM (Protospacer Adjacent Motif), that is recognized by the protein component of the complex (Semenova, Jore, Datsenko, Semenova, Westra, Wanner, et al., 2011). PAM recognition is a prerequisite for protospacer recognition, during which the crRNA base pairs with the complementary target strand of the DNA duplex. The non-target strand is hereby expelled, leading to the formation of an RNA-DNA hybrid called R-loop (Jore, Lundgren, van Duijn, Bultema, Westra, Waghmare, et al., 2011). Successful R-loop formation triggers the subsequent DNA cleavage.

A broad range of techniques have been applied to decipher the molecular mechanisms of CRIPSR-Cas systems. The combination of *in vivo* and *in vitro* studies, associated to structural snapshots, allowed the understanding the pathways of CRISPR systems as briefly described above (Fig. 1). Among the *in vitro* approaches, single-molecule tools have uniquely revealed the dynamics of the RNP complexes during protospacer recognition. Single-molecule fluorescence experiments, such as fluorescence-resonance-energy transfer (FRET) and DNA curtain assays were able to monitor protospacer binding by the Type IE surveillance complex Cascade (Blosser, Loeff, Westra, Vlot, Kunne, Sobota, et al., 2015, Redding, Sternberg, Marshall, Gibb, Bhat, Guegler, et al., 2015). Furthermore, the dynamic search of DNA targets by the type II RNP Cas9 taking place by a three-dimensional diffusion binding mechanism could be followed (Sternberg, Redding, Jinek, Greene & Doudna, 2014). As an alternative approach, our group applied a force-based technique, specifically magnetic tweezers, to study the target recognition of CRISPR-Cas systems. Compared to the fluorescence approaches, magnetic tweezers are uniquely able to monitor in real time the formation and the extent of the actual R-loop structure. Additionally, the dependence of R-loop formation on the applied mechanical stress (torque from DNA supercoiling) can be studied (Szczelkun, Tikhomirova, Sinkunas, Gasiunas, Karvelis, Pschera, et al., 2014). This technique revealed unique insight into the directionality of R-loop formation, the R-loop stability and the necessary conditions for DNA target cleavage (Rutkauskas, Sinkunas, Songailiene, Tikhomirova, Siksnys & Seidel, 2015). In this chapter, we provide an overview about magnetic tweezers investigations of CRISPR RNP complexes from type I-E (Cascade complex) and type II (Cas9 complex) and explain how different aspects of the target recognition process can be dissected using this methodology.

II. Single-molecule magnetic tweezers experiments: technical aspects

Basis of the detection of R-loop formation by CRISPR-Cas RNPs using magnetic tweezers is the unwinding of the DNA duplex during this process. The crRNA base pairs with the complementary target strand while expelling the non-complementary strand. This creates an unwinding bubble within the DNA similar to a formed transcription bubble. For a DNA molecule with constraint ends, e.g. a plasmid, the number of times that the two strands of the DNA helix wrap around each other – called linking number *LK* – is invariant if both strands remain intact (not broken). The local linking number removal at the bubble has therefore to be compensated by adding to the same extent an additional linking number change to the remainder of the plasmid DNA. For example, the 32 or 33 bp long protospacer targeted by the Cascade complex is unwound by about 3 turns upon R-loop formation. This unwinding would cause positive supercoiling of the plasmid by 3 turns upon R-loop formation, such that for a typical plasmid the negative supercoiling would be reduced by an equivalent number of turns. Partial R-loop formation, e.g. on targets with mismatches, would result in correspondingly lower supercoiling changes. Magnetic tweezers allow to monitor such supercoiling changes on single DNA molecules in real-time when appropriate conditions are applied (Strick, Allemand, Bensimon, Bensimon & Croquette, 1996). They are thus an ideal tool to study R-loop formation by CRISPR-Cas enzymes. Magnetic tweezers-based supercoiling measurements have been applied to a broad range of DNA enzymes, e.g. to study supercoil release by topoisomerases (Howan, Smith, Westblade, Joly, Grange, Zorman, et al., 2012, Strick, Croquette & Bensimon, 2000), transcription bubble formation by RNA polymerases (Revyakin, Liu, Ebright & Strick, 2006), DNA wrapping by nucleosomes (Koster, Croquette, Dekker, Shuman & Dekker, 2005, Vlijm, Lee, Lipfert, Lusser, Dekker & Dekker, 2015), and most recently R-loop formation by effector complexes of the CRISPR system (Rutkauskas et al., 2015, Szczelkun et al., 2014).

In the following the magnetic tweezers technique is described in more detail and it is explained how DNA molecules can be supercoiled with this instrument and how supercoil changes can be measured. Subsequently, it is shown how this technique can be used to study multiple aspects of target recognition by CRISPR-Cas enzymes.

II.1. The magnetic tweezers setup

The main components of a magnetic tweezers setup are an inverted microscope bearing a fluidic sample chamber (flow cell) above which a pair of magnets is mounted (Fig. 2A). The flow-cell (Fig. 2B) allows DNA or protein solutions to be flushed in or out of the field of view and also to immobilize the molecules – in our case DNA – on its surface. The key principle of molecule manipulation with magnetic tweezers is the use of superparamagnetic spherical particles (further called beads or

magnetic beads) that are attached to the surface tethered DNA molecules (Fig. 2C). The magnets generate a magnetic field gradient in the direction perpendicular to the surface, such that the bead is pulled away from the surface and a stretching force is applied to the tethered DNA molecule. The magnetic field gradient is approximately constant over the whole field of view, such that multiple molecules can be manipulated in parallel (Huhle, Klaue, Brutzer, Daldrop, Joo, Otto, et al., 2015, Ramanathan, van Aelst, Sears, Peakman, Diffin, Szczelkun, et al., 2009). For the data presented in this article we either used a home-built (Cascade measurements) or a commercial (PicoTwist, Paris, for Cas9 measurements). The described principles are identical for both setups. Differences exist only in the specific hardware components used (e.g. objective, camera, etc) and the real-time detection software (written in Labview for the home-built setup versus a commercial package for the PicoTwist apparatus). A more detailed description of the home-built setup has been recently published (Kemmerich, Kasaciunaite & Seidel, 2016).

A scheme of the setup can be seen in Figure 2A. For sample illumination, the collimated light produced by a light emitting diode (LED, 660 nm emission wavelength) passes through a lens telescope, which focuses the light on the sample plane. The transmitted light from the sample is collected by an oil immersion objective and together with a tube lens a high-magnification image of the sample is projected onto the chip of a CCD or CMOS camera. A pair of permanent NeFeB magnets within an iron holder is located just above the flow-cell. The magnet holder is connected to a motorized stage controlled by servo motors, such that the magnets can be moved vertically for force adjustment (with a precision of 10 µm) and rotated (for DNA supercoiling). The range of forces that can be produced by these magnets on superparamagnetic spherical particles with 1 µm diameter is 0.01 to 7 pN. Larger forces can be achieved by using larger magnetic beads or stronger magnets. The flow-cell is composed of two cover slips that sandwich a para-film layer in between (Fig. 2B). The channel where the sample resides is cut-out of the para-film layer and the upper glass slide contains two holes that allow access to the channel. The lower glass slide is coated with polystyrene that enables the immobilisation of anti-digoxigenin and non-magnetic reference beads. DNA molecules for magnetic tweezers experiments carry multiple biotin-modifications at one end and multiple digoxigenin modifications at the other end (see section III.1.2). The molecules are prebound at one end to streptavidin-coated magnetic beads and are then introduced into the flow cell to allow tethering of the other end (Figure 2C). The parameter that is determined in magnetic tweezers experiments is the DNA length, which is obtained from the position of the magnetic bead with respect to an immobilized reference bead. To track the positions of the magnetic and the reference bead, the beads are imaged in over-focus and recorded images are analysed in real-time by a home-developed (Huhle et al., 2015) or a commercial (see above) software (Fig. 2D). The lateral bead positions (along the x and y axes) are obtained by evaluating the centre of the diffraction pattern of

the bead through calculating a cross-correlation function with the mirrored pattern (Gosse & Croquette, 2002, Otto, Czerwinski, Gornall, Stober, Oddershede, Seidel, et al., 2010) (Fig. 2E). For tracking the position in perpendicular direction to the surface (along the z-axis) the radial intensity profile of the diffraction pattern is compared with a reference set (Gosse et al., 2002, Huhle et al., 2015). Such a reference set (or look-up table) is obtained by recording a stack of bead images at consecutively increasing over-focus positions (Fig 2F). Precise changes of the over-focus are obtained using a piezoelectric nano-positioning stage onto which the objective is mounted.

II.2. DNA stretching and force calibration

After introducing bead-bound DNA molecules into the flow cell and allowing their tethering to the bottom glass slide, unbound beads are removed from the surface by approaching the magnet to the flow cell followed by flushing the cell with buffer. Tethered DNA molecules, i.e. being attached between the magnetic bead and the surface of the flow cell, are stretched by the magnetic force that the magnetic field gradient generates on the bead. By changing the magnet position above the flow cell, the force acting on a particular bead can be precisely adjusted. To this end the force is calibrated (always for a specific magnet position) using the thermal fluctuations of the magnetic bead in lateral direction, i.e. perpendicular to the magnetic force. Force calibration is carried out individually for every magnetic bead. By considering the inverted pendulum geometry of the bead-DNA system and applying the equipartion theorem a simple relationship between force F and the mean square displacement $\langle \Delta x^2 \rangle$ of the bead in lateral direction x can be derived (Gosse et al., 2002):

$$F = \frac{k_B T}{\langle \Delta x^2 \rangle} L$$

where k_B is the Boltzmann constant, T the temperature and L the DNA length at the given force (i.e. the mean height of the bead above the surface). Note that L depends on the applied force, since DNA in absence of external tension adopts a random coil configuration, such that a stretching force is required to overcome the thermal fluctuations and bring the molecule in an extended configuration. The mean-square-displacement is determined from the frequency spectrum of the bead fluctuations in order to account for camera sampling artifacts and low frequency drift. An improved methodology together with a software implementation of the force calibration protocol has been recently published (Daldrop, Brutzer, Huhle, Kauert & Seidel, 2015).

II.3. DNA twisting and torque on DNA

The DNA molecule is attached at each end at multiple points using several digoxigenin/anti-digoxigenin linkages at the bottom of the flow-cell and several biotin/streptavidin linkages at the magnetic bead. The molecule is therefore torsionally constrained between bead and surface, which is used for DNA supercoiling (Fig. 3a). Due to a residual magnetization anisotropy of the beads (Klaue & Seidel, 2009), rotation of the magnets (i.e. the magnetic field) induces a synchronous rotation of the magnetic bead. Thus each magnet rotation induces one turn (supercoil) into the DNA. The occurring transitions during DNA twisting have been described previously in detail (Brutzer, Luzzietti, Klaue & Seidel, 2010, Mojica et al., 2009, Strick, Allemand, Bensimon & Croquette, 1998). When twisting a torsionally relaxed DNA under tension, the DNA length initially stays approximately constant (Fig. 3a). In this phase the twist, i.e. the torsion, in the DNA increases linearly with the number of added turns. When a critical torsion is reached, the DNA molecule suddenly buckles accompanied by a sudden drop of the torsion (Brutzer et al., 2010, Schopflin, Brutzer, Muller, Seidel & Wedemann, 2012) and a so-called plectoneme forms (Fig. 3a). In this structure, that is also well-known for plasmid DNA, the DNA duplex wraps around itself by forming a superhelix. Additionally added turns are absorbed topologically in form of writhe, i.e. the number of crossings that the duplex forms with itself within the superhelix. In the plectonemic phase, each added turn is thus absorbed by the linearly expanding plectoneme providing that the DNA length decreases linearly upon turning. In contrast, the torque within the molecule stays constant, since a constant energy is required to expand the superhelix. At low forces the described transitions are symmetric with respect to positive and negative supercoiling (untwisting and overtwisting of the DNA helix, respectively), such that a symmetric rotation curve is obtained (Fig. 3a). At elevated forces, higher critical torques would be achieved (see below) exceeding the denaturation torque of the DNA duplex. Thus, at these forces (about 0.7pN corresponding to a torque of 10 pN nm) the DNA length does not decrease steeply anymore for negative turns, since denaturation is preferred over plectoneme formation. Thus the rotation curves become asymmetric.

The linear conversion between added turns and DNA length in the plectonemic phase can conveniently be used to study DNA (un-)twisting by DNA enzymes. As explained above, a local DNA unwinding (e.g. by R-loop formation) reduces the applied negative supercoiling, which is thus seen as a DNA length change (see in more detail below). As long as a plectoneme remains extruded the torque in the molecule remains approximately constant during such twist changes. The torque in this post-buckling phase depends on the applied stretching force (Fig. 3b). This dependency has been characterized in detail before (Brutzer et al., 2010, Forth, Deufel, Sheinin, Daniels, Sethna & Wang, 2008, Lipfert, Kerssemakers, Jager & Dekker, 2010, Maffeo, Schopflin, Brutzer, Stehr, Aksimentiev, Wedemann, et al., 2010, Mosconi, Allemand, Bensimon & Croquette, 2009, Oberstrass, Fernandes & Bryant, 2012, Schopflin et al., 2012). For our experiments we use a force to torque conversion based

on a previously established model for which also a software tool is provided (Maffeo et al., 2010, Szczelkun et al., 2014). By studying enzyme-driven DNA untwisting on supercoiled DNA at different stretching forces allows thus to obtain insight into the torque dependence of this reaction.

III. Studying CRISPR-Cas sytems of *Streptococcus thermophilus*

The *S. thermophilus* (*St*.) DGCC7710 model organism possesses four distinct CRISPR systems, each containing a set of Cas genes and a CRISPR array, including representatives from the three major types (I to III) that have been shown to function independently (Horvath & Barrangou, 2010). With this strain the first *in vivo* evidence of direct spacer acquisition and interference activity was demonstrated for its two distinct Type II-A CRISPR systems (Barrangou et al., 2007, Garneau, Dupuis, Villion, Romero, Barrangou, Boyaval, et al., 2010). Intensive *in vitro* biochemical work carried out in the Siksnys laboratory at Vilnius University revealed the molecular mechanisms of each *St*RNP CRISPR-Cas complex starting from nucleic acid target recognition to DNA cleavage. This included the characterization of the Cascade complex and the Cas3 helicase-nuclease from the Type I-E system (Sinkunas et al., 2013, Sinkunas et al., 2011), the endonuclease Cas 9 from the Type II-A systems (Gasiunas et al., 2012), and the CSM complex from the Type III-A system (Kazlauskiene et al., 2016, Tamulaitis et al., 2014). Taking advantage of the available recombinant versions of *Streptococcus thermophilus* CRISPR RNP complexes, it was possible to carry out single-molecule characterizations of R-loop formation and DNA processing by Cas9 and Cascade/Cas3 using the magnetic tweezers technique (Rutkauskas et al., 2015, Szczelkun et al., 2014).

III.1. Monitoring real time R-loop formation by Cas9 and Cascade from *S. thermophilus*

III.1.1 General principle of R-loop detection by Cascade and Cas9

As introduced before, the basis for the detection of R-loop formation by magnetic tweezers is the local duplex unwinding at the target site. Due to the conservation of the linking number on torsionally constraint DNA, local DNA unwinding causes a positive supercoiling change of the same magnitude. Therefore on negatively supercoiled DNA the negative supercoiling will be reduced by the number of unwound helical turns of the R-loop. The supercoiling curve becomes therefore shifted to the negative side with respect to the curve recorded in absence of the R-loop (Fig. 4A). Similarly, if the molecule is kept at a fixed negative turn number, R-loop formation is seen as an abrupt length increase, since the negative supercoiling is reduced (see cartoon in blue arrow in rotation curve in Fig. 4A). If the R-loop is unstable, i.e. R-loop formation and dissociation are in dynamic equilibrium, R-loop dissociation would be seen as an abrupt DNA length decrease of

equivalent size at negative rotations. However, often R-loops of sufficient length become highly stable, such that positive torsion is required to dissociate them (see below). In this case, when recording a full rotation curve from negative to positive turns, not only the negative side but also the positive side of the rotation curve appears to be shifted. The shift at the positive side may be very stable, such that a higher positive torque (by applying a higher stretching force) is required for R-loop dissociation (Fig. 4B, see red arrow in data for Cascade). Alternatively, the R-loop becomes already unstable at lower positive torque, such that its dissociation occurs during rotation (Fig. 4C, red arrow in data for Cas9). Dissociation events at positive torque are seen as length increases, since DNA rewinding at the target site removes positive supercoiling from the DNA.

From this basic measurement scheme, i.e. that R-loop formation and dissociation can be seen as abrupt DNA length changes or as a shit of the rotation curve, a number of different parameters as function of the applied torque can be obtained. These parameters include the time to form a (stable) R-loop, the extent/length of a formed R-loop (from the observed shift of the supercoiling curve or from the size of the jump) and the stability of the formed R-loop.

III.1.2 Experimental procedures

After introducing the measurement principle, more specific experimental details are provided in the following. We also like to refer to a very detailed related protocol on how to carry out DNA supercoiling experiments within magnetic tweezers (Revyakin et al., 2006, Revyakin, Ebright & Strick, 2005).

All proteins were produced according to previously published methods by the Siksnys laboratory (Gasiunas et al., 2012, Sinkunas et al., 2013). The DNA constructs were prepared from three fragments; two 0.6 kbp long attachment handles of which one carried multiple biotin and the other multiple digoxigenin modifications and a 2.2 kbp long central fragment containing the protospacer sequence. The attachment handles were obtained by PCR in presence of biotin or digoxigenin modified dNTPs (Luzzietti, Brutzer, Klaue, Schwarz, Staroske, Clausing, et al., 2011). For the central fragment, protospacer sequences of interest were cloned into a pUC19 or a pUC18 plasmid (for StCascade or StCas9 experiments, respectively). These plasmids served as a template to generate the 2.2kbp fragment by PCR. The three PCR products (digoxigenin/biotin labelled handles and central construct part) were cleaved by restriction enzymes to generate suitable sticky ends followed by overnight ligation. The full DNA construct was subsequently purified from agarose gels by avoiding exposure to ethidium bromide and ultra-violet radiation (Luzzietti et al., 2011).

To set up magnetic tweezers experiments flow-cells were prepared: The flow cell was filled with a 50 µg/ml antidigoxigenin solution to allow binding of the antibody on the polystyrene coated surface for

at least one hour. Afterwards the surface was passivated to avoid non-specific attachments of DNA or beads by incubation with 10 mg/ml BSA for at least four hours. After removing the excess of BSA using phosphate buffer saline (PBS), reference beads (Dynospheres monosized polymer particles) were added to allow their binding on the protein-covered surface. The DNA construct was mixed with a stoichiometric excess of streptavidin coated superparamagnetic beads in PBS to allow DNA tethering at the biotinylated end. The DNA-bead mixture was diluted down to 1 pg/µl DNA and was flushed into the flow-cell. After an incubation time of approximately 5 minutes to allow tethering of the digoxigenin-modified DNA end to the bottom of the flow cell, magnets were lowered to remove unbound beads. The excess of non-bound material was washed away with PBS. Suitable beads and surface tethered DNA molecules were visually identified in the microscope images and the integrity of the chosen single DNA molecules was assessed according to the DNA length, the shape of the rotation curve and the persistence length (Revyakin et al., 2005). Forces were calibrated over a range of different magnet positions as described above.

For enzyme measurements the flow cell was equilibrated with the corresponding binding/reaction buffer. For R-loop formation experiments typically a buffer containing 20 mM Tris-HCl at pH 8.0, 150 mM NaCl and 0.1 mg/ml of BSA and for DNA degradation experiments a buffer containing 10 mM Tris-HCl at pH 7.5, 75 mM NaCl, 40 mM KCl, 7% (v/v) glycerol, 1.5 mM $MgCl_2$, 0.1 mM $NiCl_2$, 2 mM ATP, and 0.1 mg/ml BSA were used. Subsequently the CRISPR-RNP complexes were flushed into the flow cell and the formation of R-loop was studied in real time (see below). Since in the single-molecule experiments transient events on a single molecule are probed, often low amounts of protein were sufficient to produce a detectable signal. For example StCas9 was typically used at 1 nM and StCascade as was used at concentrations as low as 100 pM in flow cell volumes of less than 50 ul.

III.1.3 Data collection and analysis

For the magnetic tweezers experiments time, trajectories of the DNA length at different applied forces and/or applied turns were recorded in the corresponding setup (see above) at sampling rates of 60-300 Hz. From the trajectories, the individual events were analysed. Typically evaluated parameters were the reaction times to form or dissociate an R-loop as well as the size of the associated abrupt DNA length change, which reveals the corresponding supercoiling during the transition. Reaction times were obtained from a series of individual events (see Fig. 4D for an example). Typically such reaction times are exponentially distributed as seen e.g. in plots of the cumulative distribution of the individual R-loop formation times (Fig. 4E). To obtain a meaningful mean reaction time (being the decay time constant of the exponential) for such a broad distribution, a sufficient number of events need to be evaluated. The mean reaction time for a given torque is obtained by fitting the data to an exponential relationship (Szczelkun et al., 2014) since this way data

outliers, such as occasionally occurring unusual long times impact less the fitting result. Since such a fitting often underestimates the real statistical error, the error of the mean time is calculated as the standard error of an exponential distribution (being the error of the maximum likelihood estimation). The error is thus given by the mean reaction time divided by the square root of the number of events. Individual reaction times were extracted from the trajectories either manually or using an automatized script (written in Matlab) that splits each trajectory into the individual R-loop formation/dissociation events.

III.2. R-loop initiation, directionality and stability (PAM recognition, effect of mismatches, locking)

As explained above R-loop formation is observed as an abrupt length change of the negatively supercoiled DNA molecule or likewise a shift of the whole rotation curve (Fig. 4B,C). In the presence of a fully matching sequence and consensus PAM, supercoiling changes associated with R-loop formation (from the shift of the rotation curve) corresponded to ~2.8 turns for StCascade (Fig. 4B) and ~1.5 turns for StCas9 (Fig. 4C) (Szczelkun et al., 2014). The measured supercoiling changes were in overall agreement with the expected values of 3.0 and 1.9 turns calculated from the target lengths of 32 bp for StCascade and 20 bp for StCas, respectively. The slightly lower values resulting from the measurements are most likely due to writhe compensation caused by RNP complexes bending the DNA. Measuring the mean time for R-loop formation (Fig. 4D) at different DNA stretching, allows probing the torque dependence of this process. While R-loop formation by StCascade was strongly torque dependent – increasing negative torque caused accelerated R-loop formation (Fig. 5B) – R-loop formation by StCas9 was torque independent (Fig. 5C). The strong torque dependence for StCascade provided a rapid R-loop formation already at low negatives torques around -1.5 pN nm (Fig. 5B). R-loop dissociation was only observed at positive supercoiling and was for both complexes torque dependent (Figs. 2B and 2C in (Szczelkun et al., 2014)). While StCas9 dissociated at comparably low torques (from 4 to 8 pN nm), the StCascade bound to a fully matching target was much more stable and dissociated on experimentally relevant time scales only at torque values of around 20 pN nm. Varying the concentration of the RNP complexes affected only the R-loop formation time, but not R-loop dissociation (Fig. 2E in (Szczelkun et al., 2014)) suggesting that initial RNP complex binding to the target is the rate-limiting step during the full R-loop formation process.

By applying this characterization - revealing the kinetics of R-loop formation, the stability of the R-loop as well as the associated supercoil change - to altered target sequences the mechanism of R-loop formation can be dissected in detail. Specifically, we demonstrate the influence of the PAM

motif that is recognized by the protein itself, single mismatches in the target sequence as well as extended mismatches at the PAM distal end.

An important prerequisite for target recognition is the presence of a permissive (suitable) PAM sequence upstream of the protospacer (Leenay, Maksimchuk, Slotkowski, Agrawal, Gomaa, Briner, et al., 2016, Mojica et al., 2009). The PAM motif allows the effector complexes to distinguish spacers in the CRISPR array (non-permissive PAM) from target protospacers (with permissive PAM). For StCascade, the consensus PAM sequence (as used above) on the non-targeted strand is the dinucleotide AA for StCascade. Using a (less permissive) TT dinucleotide as PAM, the R-loop formation times are still strongly torque-dependent. However, for a given torque R-loop formation greatly slowed down or seen from a different perspective, larger negative torques are required for R-loop formation on a similar time scale (Fig. 5B). More dramatically, for observing R-loop formation with a protospacer containing a CC PAM motif (dinucleotide in CRISPR array spacer), the dsDNA needed to be virtually denatured (-200 turns at 2 pN force) to allow StCascade to bind to the target sequence (Fig. 3A in (Szczelkun et al., 2014)). Similarly, also R-loop formation by StCas9 was strongly PAM dependent. A single nucleotide substitution in the consensus PAM motif NGGNG (Gasiunas et al., 2012) provided more than an order of magnitude increase in the mean R-loop formation time (Fig. 5C). In contrast to the strong influence of the PAM on R-loop formation, the R-loop stability was not affected for both types of CRISPR systems. This data allows concluding that PAM recognition is important during the initiation of R-loop formation (Fig. 5A) (Szczelkun et al., 2014).

By introducing a single mismatch between the target DNA sequence and the crRNA of StCascade, it could be directly shown that R-loop formation employs a directional zipping mechanism, i.e. it starts the PAM motif and moves towards the PAM distal end of the protospacer (Fig. 5D). The introduction of a single mismatch leads to the formation of an intermediate R-loop corresponding to DNA unwinding between the PAM motif to the mismatch. This partial R-loop either collapses (dissociation of the RNP complex) or bypasses the mismatch to subsequently form the full R-loop (Fig. 5E). By moving the mismatch further from the PAM, the observed intermediate R-loops became longer, pin-pointing the partial R-loops to the region between PAM and mismatch in agreement with directional zipping (Fig. 5F).

While intermediate R-loops for StCascade are highly unstable and readily dissociate even at negative torque (Fig. 5E), a high positive torque is needed to remove full R-loops from the DNA (Fig. 4B). To test the sequence requirements for this stability change, mismatches between the DNA target and the crRNA at the PAM distal end can be employed. Introducing 6 (or more) base-pair mismatches at the PAM distal end caused all R-loops to become unstable (Fig. 5H), such that they collapse instantaneously as soon positive supercoiling is applied. Introduction of a 4 base-pair mismatch

caused more than 50% of the R-loops to become unstable, while the remaining R-loops where highly stable comparable to the fully matching target (Fig 5H). Even for a 2 bp PAM-distal mismatch a fraction of the R-loops was unstable. The existence of two R-loop stabilities for a given substrate suggests a conformational rearrangement that "locks" an R-loop once it reached the PAM distal end (Fig. 5G). The observed "locking" is consistent with two structural snapshots of *E. coli* Cascade made by cryo-electron microscopy in presence or absence of an R-loop (Wiedenheft, Lander, Zhou, Jore, Brouns, van der Oost, et al., 2011). Thus, the locking transition acts as a sensor for full R-loop formation.

When investigating R-loops from StCas9 on substrates with mutations over the PAM distal end, a somewhat different picture was obtained. Such mutations did not change the R-loop formation time (Szczelkun et al., 2014). The R-loop stability (i.e. the R-loop dissociation time at positive supercoiling), being lower than for StCascade was only moderately affected by up to 7 PAM-distal mismatches (Fig. 5I). No R-loop formation was however seen when introducing 9 or more PAM-distal mismatches. Therefore, for StCas9 R-loops of just 13 bp length are already considerable stable, while shorter R-loops are highly unstable and can even not be introduced by considerable negative torque. The R-loop stabilization already at a moderate length is in agreement with recent crystallographic data of Cas9 bound and not bound to target DNA. Comparison of these structures suggested a conformational change once the R-loop zipped to a length of 10 bp (Anders & Jinek, 2014, Jiang, Zhou, Ma, Gressel & Doudna, 2015, Jinek, Jiang, Taylor, Sternberg, Kaya, Ma, et al., 2014).

III.3. Monitoring DNA cleavage using multi-molecule experiments

As illustrated above the observed locking transition for StCascade is suggested to act as a sensor for full R-loop zipping. This sensor is supposed to control the downstream recruitment of Cas3 (Hochstrasser, Taylor, Bhat, Guegler, Sternberg, Nogales, et al., 2014) and thus the DNA degradation. Cas3 first displays an endonuclease activity by nicking the DNA on the non-target strand (NTS) of the protospacer and then further degrades DNA upstream of the protospacer in the 3' to 5' direction of the NTS (Sinkunas et al., 2013, Westra, van Erp, Kunne, Wong, Staals, Seegers, et al., 2012). To directly correlate locking with DNA degradation and to test further sequence requirements one can carry out DNA degradation experiments, in which first R-loops are introduced on a substrate of choice in a highly controlled manner and subsequently DNA degradation is monitored by addition of Cas3 (Rutkauskas et al., 2015).

As the DNA cleavage ultimately produces a dsDNA break, the DNA-bead tether will be lost. Since DNA cleavage is a single disruptive event that cannot be repeated on a given molecule, a magnetic tweezers setup was applied for this type of measurements that supports tracking of multiple beads

at the same time (Huhle et al., 2015, van Aelst, Toth, Ramanathan, Schwarz, Seidel & Szczelkun, 2010). Compared to conventional magnetic tweezers as used above, such a setup uses a reduced magnification of the microscope as well as a camera with a larger pixel array (see (Kemmerich et al., 2016) for details).

To carry out Cas3 dependent DNA degradation experiments, StCascade was added into the flow-cell and R-loop formation was assisted by negative supercoiling. After providing enough time for R-loop formation at the given DNA target, R-loop formation was verified as described above. Subsequently, StCas3 in the corresponding reaction buffer (containing ATP, Mg^{2+} and Ni^{2+}) was added to the reaction (Fig. 6A), while keeping the molecules in a negatively supercoiled state. In time trajectories of the DNA length, DNA nicking, seen as a DNA length increase due to supercoil release, followed by DNA cleavage, seen as a complete loss of the magnetic bead could be observed (Fig. 6B). From cumulative distributions of the times after which cleavage of individual molecules occurred, the cleavage kinetics was obtained (Fig. 6C). Using this cleavage assay, it was uniquely possible to distinguish between R-loop formation and the actual cleavage reaction during the whole DNA degradation process. This allows to study the influence of target mutations on the Cas3-dependent cleavage step, in particular PAM alterations, internal mismatches and mutations along the PAM distal end that impair locking.

For the non-permissive CC-PAM (instead of the consensus AA dinucleotide) stable, i.e. locked, R-loops can be introduced using extensive negative supercoiling (see above). Interestingly, StCas3 cleavage was strongly impaired on R-loops containing a CC PAM sequence (Fig. 6C green triangles) compared to the consensus PAM. Thus, a permissive PAM is a prerequisite for cleavage even in the presence of a full R-loop.

On targets containing single internal mismatches near the PAM, R-loop formation has been shown to be dramatically impaired (Semenova et al., 2011) such that much higher negative torques are required (Rutkauskas et al., 2015). To test whether internal mismatches can affect cleavage, DNA degradation experiments were carried out on a target protospacer with a mismatch at position 2 from the PAM. On preformed R-loops, DNA cleavage occurred with a similar kinetics as for the fully matching protospacer with the consensus PAM (Fig. 6C, light blue squares) suggesting that Cas3 recruitment and activity are not influenced by internal mismatches. A similar conclusion was obtained for a mismatch at position 22 from the PAM (Rutkauskas et al., 2015).

To unambiguously test the influence of locking, DNA degradation of a target bearing 8 mismatches at the PAM distal end was tested, where locking is abolished. Despite keeping the (unlocked) R-loop

introduced with the help of negative supercoiling, cleavage was completely abolished thus demonstrating the need for a locking state to recruit Cas3 (Fig. 6C black diamonds).

These experiments reveal that there are two prerequisites for DNA degradation by StCas3 – an R-loop that zipped fully down to the PAM distal end of the protospacer and simultaneously matching PAM. However, mismatches that are encountered and efficiently bypassed during R-loop formation still permit degradation by Cas3 with similar rate. Thus, the successful R-loop propagation represents the actual target verification mechanism, while the CRISPR-Cas RNP just initiates the R-loop formation at the PAM and then verifies its formation.

IV. Studying E.coli cascade

The magnetic tweezers assay can be readily applied to other DNA targeting CRISPR-Cas systems in order to study their diversity. For illustration, first data of our assay are presented for the Type IE CRISPR-Cas system from *Escherichia coli*. Whereas some similar general mechanisms have been found with StCascade, some noticeable differences can also be pointed out. At a protein level, both RNP Cascade complexes have the same subunit composition even if the proteins differ in their specific sequences and length of peptide chains. Here a selected overview of results for EcCascade is presented. A more detailed study will be published elsewhere.

IV.1. Similarities and differences compared to StCascade

In *E. coli*, the common PAM sequences for interference are the trinucleotides AAG or ATG on the non-target strand of the protospacer with AAG displaying highest efficiency in immunity (Fineran et al., 2014, Hayes, Xiao, Ding, van Erp, Rajashankar, Bailey, et al., 2016, Westra, Semenova, Datsenko, Jackson, Wiedenheft, Severinov, et al., 2013). In accordance with this observation, the mean time for R-loop formation measured within magnetic tweezers is considerably lower in presence of an AAG PAM compared to an ATG PAM on a matching protospacer (unpublished data). As observed for StCascade, locking is observed once a full R-loop forms, i.e. the R-loop becomes highly stable confirming a general molecular mechanism for Type IE CRISPR-Cas systems. However, in contrast to StCascade the *E. coli* complex remained tightly bound to the DNA and only elongated time periods combined with extreme positive torques that almost denature DNA would induce its dissociation (unpublished data). Because of this, and to be able to generate statistics, most experiments have been done with protospacers having 6 mismatches with the crRNA at the PAM distal end that in agreement with StCascade abolish locking. Working in similar conditions as for StCascade, it appeared that considerably higher EcCascade concentrations are needed to obtain similar R-loop formation times (10nM of EcCascade versus 100 pM of StCascade). Moreover, in contrast to

StCascade no strong torque dependence is observed with EcCascade (Fig. 7A), although a critical minimum negative torque is needed to trigger R-loop formation, since no R-loops form at positive torque.

IV.2. Testing of reaction conditions

The relatively slow kinetics of R-loop formation for EcCascade, even at high protein concentration, can be changed by altered reaction conditions, namely the presence of Mg2+ ions and by an increase in the temperature of the flow cell (Fig. 7B). Different ionic conditions can be readily applied within the reaction buffer. Sample heating can be conveniently implemented using commercial objective heating systems. These systems need to be correspondingly calibrated to compensate for the temperature drop between the heated objective and the actual flow cell (Kemmerich et al., 2016). For the heated system it is advisable to carry out the force calibration at room temperature prior applying the higher temperature due to a significantly altered viscosity of the aqueous medium (Venable, Hatcher, Guvench, Mackerell & Pastor, 2010).

Addition of 5 mM Mg^{2+} increased the R-loop formation rate about 2-fold (Fig. 7B). A further increase of the formation rate was obtained at 37 ˚C summing up to about an order of magnitude in comparison to the commonly used conditions. The influence of the Mg^{2+} ions may be due to their ability to stabilize secondary structures of DNA such as hairpins. Therefore it may facilitate the opening of the DNA strands during R-loop initiation. Alternatively, the protein-nucleic acid interaction may be altered.

V. Perspectives and conclusion

V.1. Unified view on target recognition and verification by Cascade and Cas9

A detailed mechanistic understanding was the driving force to employ CRISPR-Cas effector complexes, especially Cas9, as unique tools in genome editing and gene silencing applications. The simplicity by which a gene editing tool for a desired gene can be generated surpasses previously existing programmable gene targeting tools such as rare-cutting Zn-finger nucleases and TALENs by far. However, CRISPR-Cas systems also recognise and degrade DNA targets containing multiple mismatches – a property that is certainly beneficial in phage defence to avoid phage escape by mutation. This causes off-targeting and is highly undesirable in biotechnology applications. Multiple top-down approaches to evaluate the effects of off-targeting by Cas9 showed highly promiscuous recognition (Kuscu, Arslan, Singh, Thorpe & Adli, 2014, Hsu, Scott, Weinstein, Ran, Konermann, Agarwala, et al., 2013, Pattanayak, Lin, Guilinger, Ma, Doudna & Liu, 2013). The magnetic tweezers experiments presented here as well as other biochemical and single-molecule investigations, now

provide a unified mechanism how DNA targets are recognized by Type I and Type II CRISPR-Cas systems (Fig. 8A):

Both, Cascade and Cas9 search for PAMs through 3D diffusion (Redding et al., 2015, Sternberg et al., 2014) and after PAM binding R-loop formation is initiated (Szczelkun et al., 2014). R-loops expand from the PAM by zipping along the target strand. If a mismatch is encountered R-loop formation stalls (Rutkauskas et al., 2015), since this represents an energetic barrier for further propagation. Such an R-loop intermediate can either overcome the mismatch and continue zipping or collapse. In the latter case the RNP would leave the target and reinitiate the search for a better matching target (supported by the strong concentration dependence of R-loop formation). After full R-loop zipping up to the PAM-distal end, a conformational changes is triggered for both, Cascade and Cas9, that in turn promotes DNA degradation (Wiedenheft et al., 2011, Szczelkun et al., 2014). Thus, the actual target recognition mechanism is a passive zipping between two nucleic acid strands. In this simplified view, the sole task of the protein complex is to effectively initiate the R-loop by a PAM dependent interaction and to verify whether the full zipping process was successful. Both, RNP complex types differ however regarding the verification mechanism of full zipping. For Cascade full zipping “locks” the R-loop (Szczelkun 2014), i.e. it becomes hyperstable, due to a large conformational change (Wiedenheft et al., 2011). Locking is the prerequisit for Cas3 recruitment and subsequent DNA degradation and DNA cleavage itself is not affected by mismatches between target DNA and crRNA (Hochstrasser et al., 2014, Rutkauskas et al., 2015). In case of Cas9, R-loop gain already considerable stability once zipping over just a bit more than half the target length (Szczelkun et al., 2014). For such stable but not fully zipped R-loops, DNA nicking and cleavage is however strongly impaired (Szczelkun et al., 2014). DNA cleavage requires the engagement of the HD-nuclease domain. This conformational change occurs however only once the R-loop zips until the PAM distal end (Sternberg et al., 2014)providing a high similarity to R-loop verification by Cascade. Overall, a detailed mechanistic understanding of the zipping process and the subsequent target verification would provide the basis to quantitatively predict possible off-target effects in order to ialready existing genome editing and gene silencing tools.

V.2. Challenges in single-molecule studies of CRISPR-Cas systems

Single-molecule experiments made an important contribution to understand the mechanism of DNA targeting by CRISPR-Cas systems. The experiments published so far can be grouped into different categories: (i) localization experiments on stretched DNA, in particular DNA curtains (Sternberg et al., 2014, Redding et al., 2015), (ii) Förster Resonance Energy Transfer (FRET) experiments (Blosser et al.,

2015) and (iii) the magnetic tweezers experiments presented in this manuscript. Localization and FRET experiments probed the binding behaviour of Cascade and Cas9 on DNA targets with different sequences, elucidating e.g. the scanning of the PAM sequence by these enzymes. Localization experiments were in addition used to observe the recruitment of additional proteins, in particular Cas3, and their movement along DNA that provided detailed insight into Cas3-based DNA degradation as well as primed adaptation. Alternatively magnetic tweezers probe directly the R-loop formation and can uniquely measure the extent of formed R-loops as well as clearly distinguish and assign intermediate states.

A future challenge in mechanistic studies of CRISPR-Cas systems will be to combine the information of the different experimental approaches, for example to probe the conformational changes of the involved enzymes (e.g. locking) and to correlate them with downstream events (e.g. Cas3 recruitment). A first step to gain insight into conformational changes of these enzymes has been taken in a FRET study (using bulk FRET) where it was shown that upon full R-loop formation the HNH nuclease domain of Cas9 gets engaged (Sternberg). Engagement of this domain was the prerequisite for DNA cleavage.

In order to combine different information/approaches, we constructed a setup that integrates high-resolution magnetic tweezers and dual-colour single-molecule fluorescence detection into a single instrument (Kemmerich, Swoboda, Kauert, Grieb, Hahn, Schwarz, et al., 2016, Brutzer, Schwarz & Seidel, 2012). Probing the opening and closing transitions of a DNA hairpin under tension demonstrated the synchronous detection of this transition in both channels in parallel. Such an approach would also be useful to gain more information on the molecular mechanism of CRISPR-Cas RNP complexes. Fluorescent labels can be placed on the DNA or on the protein depending of the question to answer. For example, using fluorescently labelled proteins in such experiments would provide insight into the sequential processes of PAM binding, R-loop formation, Cas3 recruitment and DNA degradation (Fig. 8B). It will be particularly useful to see in that context the influence of the PAM sequence as well as the influence of mismatches between the DNA and the crRNA, e.g. whether the target scanning in form of the formation of instable R-loop intermediates is linked to multiple associations-dissociations of Cascade (Fig. 5E). Alternatively, FRET labels within a CRISPR-Cas complex will allow monitoring the dynamics and the timing of molecular rearrangement upon full R-loop formation. It will be of interest, in the case of StCascade, to elucidate the molecular rearrangement taking place as consequence of the formed R-loop and the PAM sequence used. For example, with a CC PAM sequence the full generated R-loop is locked but Cas3 is not recruited (Fig. 6C) suggesting a partial conformational change in comparison to an R-loop lacking 6 base-pairings at the PAM distal end where no rearrangement takes place (no locked R-loop, no Cas3 recruitment).

Acknowledgements

We would like to thank Virginijus Siksnys, Tomas Sinkunas, Giedrius Gasiunas and Inga Songailiene from Vilnius University for providing recombinant proteins from *Streptococcus thermophilus*. The *E. coli* Cascade expression construct and the purification protocol have been kindly provided by Konstantin Kuznedelov and Konstantin Severinov from the Waksman Institute of Microbiology, Rutgers University. We would also like to thank Dominik Kauert for valuable technical assistance for the magnetic tweezers measurements.

Figure Legends

Figure 1. General model for CRISPR/Cas systems. **A.** Schematic view of a prokaryotic cell getting infected by a virus. In the adaptation stage, the protein complex Cas1-Cas2 (in blue) is capturing a short piece of DNA virus (in orange) which is then integrated in a CRISPR locus on its genome. The newly acquired sequence (called spacer) is integrated in-between repeat sequences (in black) just beside a leader sequence containing a promoter (in yellow). At this stage, the cell and its progeny will be immunised against the virus. The RNA CRISPR locus transcript is specifically cleaved by a ribonuclease in the repeat sequences generating small RNAs (called crRNAs), each carrying a memory of viral sequence. A single crRNA is carried by a large protein or protein complex that will search for the complementary sequence in any cellular DNA. The recognition of a matching target (called protospacer) will lead to the degradation of the latter, either by the recruitment of an accessory nuclease or by the intrinsic activity of the RNP complex. **B.** In the vicinity of the CRISPR locus are found Cas genes coding for the proteins implicated in the molecular process. Bioinformatics studies allowed the classification of different types and subtypes of CRISPR systems. Whereas in all systems are found the adaptation proteins Cas1 and Cas2, the effector complexes involved in interference are quite different in composition. The type II is the large monomer Cas9 and the types I and III are multimeric with a common Cas7 protein forming a backbone on the crRNA. For the Type I, II and III are represented the subunits of the main studied CRISPR RNP complexes (Cascade, Csy, Cas9, CMR and CSM) with their respective protein signatures being Cas3, Cas9 and Cas10. **C.** Schematic representation of the effector complexes from 3 types. In contrast to the types III targeting RNA with their crRNA, the type I and II are targeting DNA and are generating an R-loop which will trigger the cleavage. The Cascade and Csy complexes recruit the fused helicase nuclease Cas3 whereas Cas9 induces a double strand break by dual incision.

Figure 2. Magnetic tweezers setup and basic principles of bead tracking. **A.** Simplified model of magnetic tweezers setup. The setup consists of LED light source, a pair of movable permanent magnets, a flow-cell, an objective and a camera. Light from a LED is being directed onto the sample which resides inside a flow-cell. After passing the sample the light is focused by an objective acting as an inverted microscope, towards the camera where the view of the sample is imaged. **B.** The flow-cell consists of two glass slides sandwiched with a thin layer of para-film. Two holes in the upper glass slide allow the sample to be pumped in or out of the channel that is formed by para-film in between the slides. **C.** The surface of the bottom glass slide inside the channel is covered by magnetic and “reference” beads. Reference beads are tightly attached on the surface of the flow-cell, while magnetic beads are tethered to the bottom glass slide by single DNA molecule. This is achieved by modifying DNA ends with digoxigenin and biotin, which binds to anti-digoxigenin and streptavidin respectively. **D.** Beads in off-focus produce diffraction rings. Diffraction rings pattern in over-focus is used for bead tracking. **E.** The light intensity profile of the bead in over-focus is used to determine the center of the bead and to track the movement of the bead in x and y-axes. **F.** Collection of diffraction pattern images in multiple over-focus positions is used to track the bead movement in z-axis.

Figure 3. Scheme of single DNA molecule supercoiling in the magnetic tweezers and relationship between twist, force and torque. **A.** Force acting on the magnetic bead and thus changing the distance between the magnets and the bead can control single DNA molecule. The pair of magnets can also be rotated thus introducing negative or positive supercoiling on the DNA molecule. Introduction of negative or positive supercoiling on the DNA causes a plectoneme formation that changes the position of the bead, which is observed as a typical "bell-shaped" rotation curve. **B.** Twisting the DNA molecule in the extended state (no plectonemes) accumulates twist leading to a torque increasing linearly on the DNA. After the first plectoneme formation, the torque on the DNA remains constant independently of additional turns applied, whereas the DNA length decreases linearly. Figures adapted from (Maffeo et al., 2010, Szczelkun et al., 2014).

Figure 4. Detection and quantification of the single R-loops formed by StCascade and StCas9 complexes. **A.** The scheme of the R-loop induced change of DNA supercoiling. Blue rotation curve represents the DNA supercoiling in absence of StCascade/StCas9, green – in presence of the formed R-loop. R-loop formation (blue arrow) may be observed in two different ways: (i) as an abrupt change of the supercoiling of the DNA or (ii) as a shift of rotation curve. Positive torque induced on the DNA causes R-loop dissociation (red arrow). **B.** The left panel shows the time trace of one Cascade R-loop cycle. Blue line represents supercoiling DNA from positive to negative turns and waiting time for an R-loop to form. The green line represents the R-loop formed and the supercoiling to positive turns. The red line represents the increased force phase including R-loop dissociation. The light blue trace on the background is the original trace that has not been smoothed. On the right panel DNA length is plotted as a function of turns applied on the DNA molecule. Colors are the same as on the left panel; the light grey line represents the rotation curve without bound StCascade. **C.** The R-loop cycle of StCas9. Colors are indicated as in (B). R-loop by StCas9 forms during rotation to negative turns (blue arrow) and dissociates during rotation to positive turns (red arrow). While StCas9 dissociates at low positive torque, increased forces are not applied. **D.** The time trace of multiple recorded R-loop formation-dissociation events by StCascade. **E.** Fitting individual R-loop formation times showed in (D) (T_1, T_2…,T_n) using exponential fit gives mean R-loop formation time at the given torque. Adapted from (Rutkauskas et al., 2015, Szczelkun et al., 2014).

Figure 5. R-loop formation and stability. **A.** To form an R-loop, Cascade or Cas9 first recognises the essential PAM motif shown in yellow, respectively 2 nucleotides for StCascade and 5 nucleotides for StCas9. **B.** The sequence of the PAM motif plays a significant role on the time for R-loop formation. On this graph are shown the torque dependence for StCascade R-loop formation where the PAM is either AA (dark blue) or TT (light blue) on the non-targeting strand. Strong torque dependence is detected for both experiments. However, a significant shift is observed between AA and TT PAM; more than a tenfold difference considering the time for R-loop formation at similar torque. **C.** The same experiment done with StCas9 also shows a tenfold difference on the time for R-loop formation considering 2 different PAM sequences. However, no significant torque dependence is observed for the complex to trigger the DNA opening. **D.** Introducing a mismatch between StCascade crRNA and the targeted strand has a noticeable effect on the R-loop stability. When the mismatch is encountered, the complex may dissociate or bypass this mismatch. **E.** A time-trace of R-loop forming on the target bearing a mismatch at the position 15 between the protospacer and the crRNA shows

events of association and dissociation of the complex before the full R-loop is formed. **F.** By moving the mismatch at positions more distant to the PAM motif, the observed intermediates have higher amplitude (light blue) suggesting R-loop directionality. Once the mismatch has been bypassed, the same full R-loop is formed (dark blue). **G.** Upon full R-loop formation, a structural conformational change of StCascade locks the complex on DNA. Inserting mismatches patch at the PAM distal-end is decreasing the stability of the complex at positive torque. **H.** Graph showing the stability of the complex at positive torque with an increasing number of mismatches at the PAM distal end. From 6 mismatches, the complex dissociates instantaneously, even at low positive torque. **I.** Comparison of the stability of StCas9 R-loops at positive torque with 7, 5, 1 or no mismatches (MM) at the PAM distal end. Adapted from (Rutkauskas et al., 2015, Szczelkun et al., 2014).

Figure 6. Cleavage by StCas3 of preformed R-loops. **A.** Scheme of StCas3 cleavage assay on tweezers. **B.** Typical time trace of single StCas3 cleavage reaction. After addition of StCas3, nicking of the DNA is observed as an abrupt jump of the bead caused by the loss of supercoiling. Disappearance of the magnetic bead indicates double-stranded DNA cleavage. **C.** Plot comparing times of individual cleavage reactions of R-loops formed on DNA bearing AA PAM, CC PAM, mismatch at the position 2 (M2) and 8 mismatches patch at the PAM distal end (8 MM). Adapted from (Rutkauskas et al., 2015).

Figure 7. Study of R-loop formation by EcCascade. **A.** Mean time for R-loop formation as function of torque, **B.** Effects of Mg^{2+} and temperature on R-loop formation: individual time points and exponential fits, **C.** like B presented in mean times.

Figure 8. Conclusion and perspectives. **A.** Unified target recognition model for CRISPR RNP complexes. (i) CRISPR RNP complex recognizes PAM and (ii) R-loop formation is initiated. (iii) R-loops zips directionally away from PAM and (iv) upon reaching PAM distal end conformational changes of the complex take place that verify the sequence (v) to destine it for cleavage/degradation. **B.** Perspective of combined FRET – magnetic tweezers experiments. Cascade or Cas9 is labeled with acceptor fluorophore, DNA with donor fluorophore to disentangle individual rates of PAM binding and R-loop formation. **C.** Two fluorophores are located on different Cascade or Cas9 subunits/domains to observe the conformational changes that are taking place during R-loop formation.

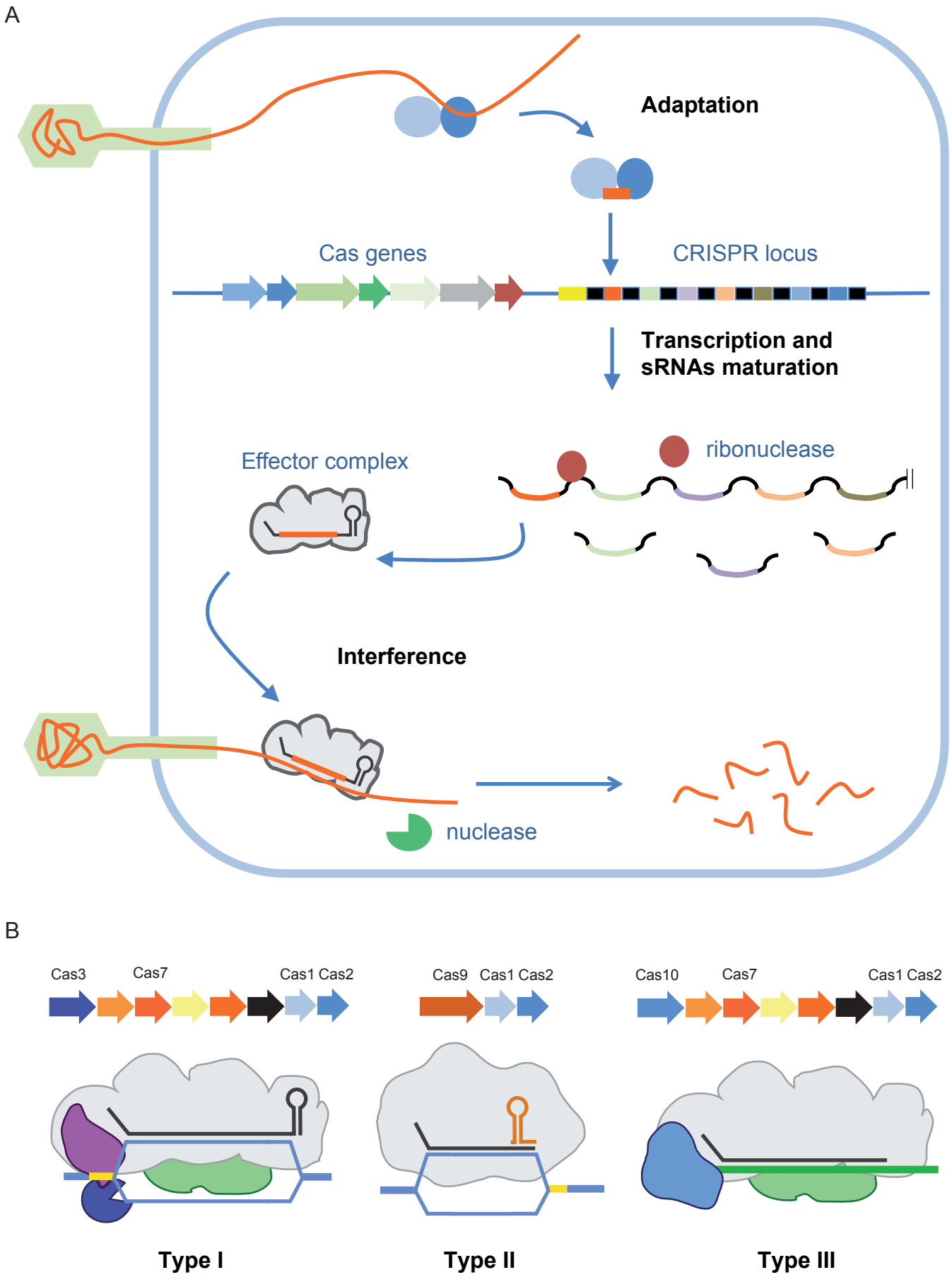
References:

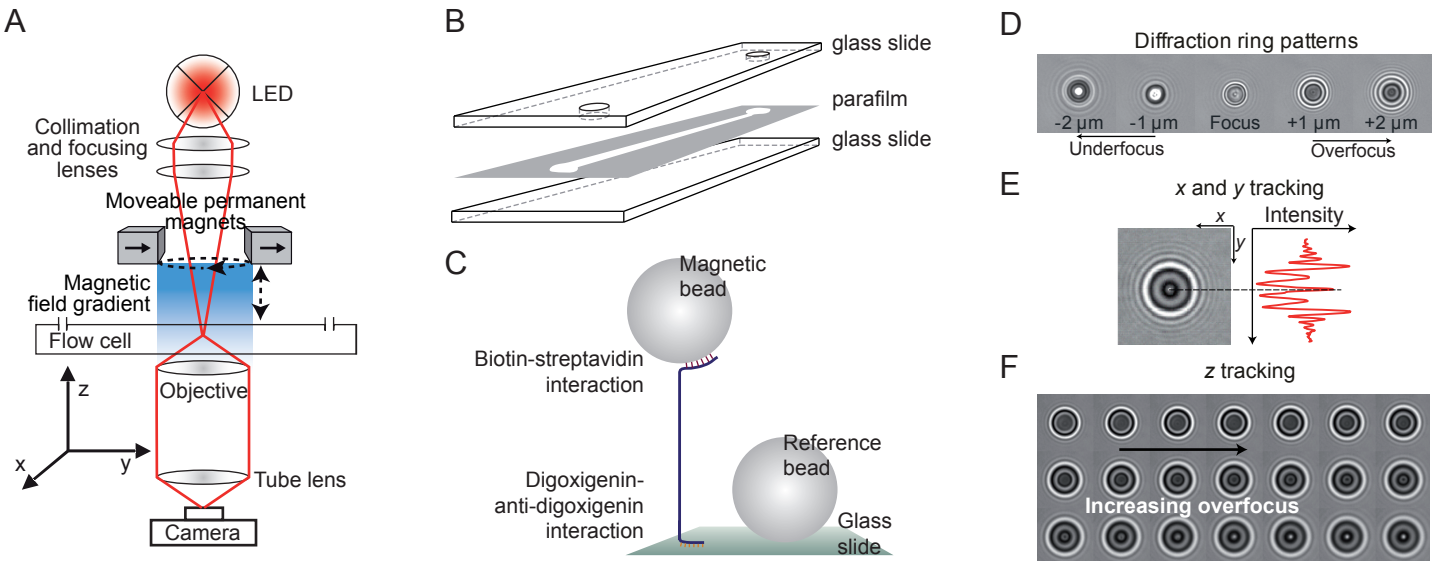
Anders, C. & Jinek, M. (2014). In vitro enzymology of Cas9. 546, 1-20.
Barrangou, R., Fremaux, C., Deveau, H., Richards, M., Boyaval, P., Moineau, S., et al. (2007). CRISPR provides acquired resistance against viruses in prokaryotes. 315, 1709-12.
Benda, C., Ebert, J., Scheltema, R. A., Schiller, H. B., Baumgartner, M., Bonneau, F., et al. (2014). Structural model of a CRISPR RNA-silencing complex reveals the RNA-target cleavage activity in Cmr4. 56, 43-54.
Blosser, T. R., Loeff, L., Westra, E. R., Vlot, M., Kunne, T., Sobota, M., et al. (2015). Two distinct DNA binding modes guide dual roles of a CRISPR-Cas protein complex. 58, 60-70.
Bolotin, A., Quinquis, B., Sorokin, A. & Ehrlich, S. D. (2005). Clustered regularly interspaced short palindrome repeats (CRISPRs) have spacers of extrachromosomal origin. 151, 2551-61.
Brutzer, H., Schwarz, F. W. & Seidel, R. (2012). Scanning evanescent fields using a pointlike light source and a nanomechanical DNA gear. 12, 473-8.
Brutzer, H., Luzzietti, N., Klaue, D. & Seidel, R. (2010). Energetics at the DNA supercoiling transition. 98, 1267-76.
Burstein, D., Sun, C. L., Brown, C. T., Sharon, I., Anantharaman, K., Probst, A. J., et al. (2016). Major bacterial lineages are essentially devoid of CRISPR-Cas viral defence systems. 7, 10613.
Charpentier, E., Richter, H., van der Oost, J. & White, M. F. (2015). Biogenesis pathways of RNA guides in archaeal and bacterial CRISPR-Cas adaptive immunity. 39, 428-41.
Daldrop, P., Brutzer, H., Huhle, A., Kauert, D. J. & Seidel, R. (2015). Extending the range for force calibration in magnetic tweezers. 108, 2550-61.
Elmore, J. R., Sheppard, N. F., Ramia, N., Deighan, T., Li, H., Terns, R. M., et al. (2016). Bipartite recognition of target RNAs activates DNA cleavage by the Type III-B CRISPR-Cas system. 30, 447-59.
Estrella, M. A., Kuo, F. T. & Bailey, S. (2016). RNA-activated DNA cleavage by the Type III-B CRISPR-Cas effector complex. 30, 460-70.
Fineran, P. C., Gerritzen, M. J., Suarez-Diez, M., Kunne, T., Boekhorst, J., van Hijum, S. A., et al. (2014). Degenerate target sites mediate rapid primed CRISPR adaptation. 111, E1629-38.
Forth, S., Deufel, C., Sheinin, M. Y., Daniels, B., Sethna, J. P. & Wang, M. D. (2008). Abrupt buckling transition observed during the plectoneme formation of individual DNA molecules. 100, 148301.
Garneau, J. E., Dupuis, M. E., Villion, M., Romero, D. A., Barrangou, R., Boyaval, P., et al. (2010). The CRISPR/Cas bacterial immune system cleaves bacteriophage and plasmid DNA. 468, 67-71.
Gasiunas, G., Barrangou, R., Horvath, P. & Siksnys, V. (2012). Cas9-crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria. 109, E2579-86.
Gilles, A. F. & Averof, M. (2014). Functional genetics for all: engineered nucleases, CRISPR and the gene editing revolution. 5, 43.
Gosse, C. & Croquette, V. (2002). Magnetic tweezers: micromanipulation and force measurement at the molecular level. 82, 3314-29.
Grissa, I., Vergnaud, G. & Pourcel, C. (2007). CRISPRFinder: a web tool to identify clustered regularly interspaced short palindromic repeats. 35, W52-7.
Hale, C. R., Zhao, P., Olson, S., Duff, M. O., Graveley, B. R., Wells, L., et al. (2009). RNA-guided RNA cleavage by a CRISPR RNA-Cas protein complex. 139, 945-56.
Hayes, R. P., Xiao, Y., Ding, F., van Erp, P. B., Rajashankar, K., Bailey, S., et al. (2016). Structural basis for promiscuous PAM recognition in type I-E Cascade from E. coli. 530, 499-503.
Hochstrasser, M. L., Taylor, D. W., Bhat, P., Guegler, C. K., Sternberg, S. H., Nogales, E., et al. (2014). CasA mediates Cas3-catalyzed target degradation during CRISPR RNA-guided interference. 111, 6618-23.
Horvath, P. & Barrangou, R. (2010). CRISPR/Cas, the immune system of bacteria and archaea. 327, 167-70.
Howan, K., Smith, A. J., Westblade, L. F., Joly, N., Grange, W., Zorman, S., et al. (2012). Initiation of transcription-coupled repair characterized at single-molecule resolution. 490, 431-4.

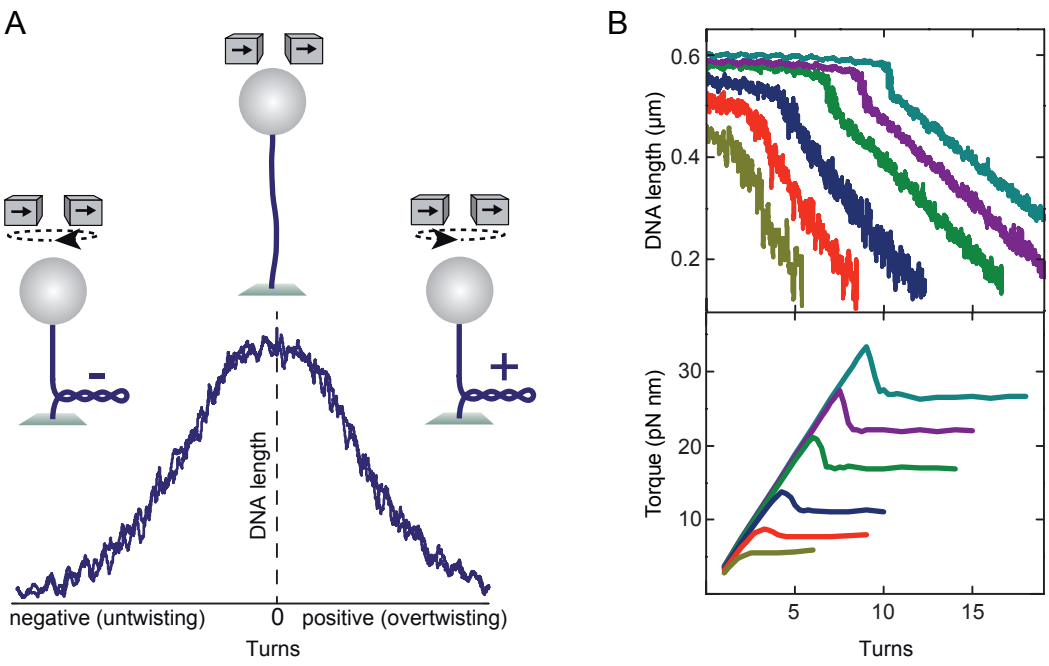
Hsu, P. D., Lander, E. S. & Zhang, F. (2014). Development and applications of CRISPR-Cas9 for genome engineering. 157, 1262-78.
Hsu, P. D., Scott, D. A., Weinstein, J. A., Ran, F. A., Konermann, S., Agarwala, V., et al. (2013). DNA targeting specificity of RNA-guided Cas9 nucleases. 31, 827-32.
Huhle, A., Klaue, D., Brutzer, H., Daldrop, P., Joo, S., Otto, O., et al. (2015). Camera-based three-dimensional real-time particle tracking at kHz rates and Angstrom accuracy. 6, 5885.
Ishino, Y., Shinagawa, H., Makino, K., Amemura, M. & Nakata, A. (1987). Nucleotide sequence of the iap gene, responsible for alkaline phosphatase isozyme conversion in Escherichia coli, and identification of the gene product. 169, 5429-33.
Jiang, F., Zhou, K., Ma, L., Gressel, S. & Doudna, J. A. (2015). STRUCTURAL BIOLOGY. A Cas9-guide RNA complex preorganized for target DNA recognition. 348, 1477-81.
Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J. A. & Charpentier, E. (2012). A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. 337, 816-21.
Jinek, M., Jiang, F., Taylor, D. W., Sternberg, S. H., Kaya, E., Ma, E., et al. (2014). Structures of Cas9 endonucleases reveal RNA-mediated conformational activation. 343, 1247997.
Jore, M. M., Lundgren, M., van Duijn, E., Bultema, J. B., Westra, E. R., Waghmare, S. P., et al. (2011). Structural basis for CRISPR RNA-guided DNA recognition by Cascade. 18, 529-36.
Karvelis, T., Gasiunas, G., Miksys, A., Barrangou, R., Horvath, P. & Siksnys, V. (2013). crRNA and tracrRNA guide Cas9-mediated DNA interference in Streptococcus thermophilus. 10, 841-51.
Kazlauskiene, M., Tamulaitis, G., Kostiuk, G., Venclovas, C. & Siksnys, V. (2016). Spatiotemporal Control of Type III-A CRISPR-Cas Immunity: Coupling DNA Degradation with the Target RNA Recognition. 62, 295-306.
Kemmerich, F. E., Kasaciunaite, K. & Seidel, R. (2016). Modular magnetic tweezers for single-molecule characterizations of helicases.
Kemmerich, F. E., Swoboda, M., Kauert, D. J., Grieb, M. S., Hahn, S., Schwarz, F. W., et al. (2016). Simultaneous Single-Molecule Force and Fluorescence Sampling of DNA Nanostructure Conformations Using Magnetic Tweezers. 16, 381-6.
Klaue, D. & Seidel, R. (2009). Torsional stiffness of single superparamagnetic microspheres in an external magnetic field. 102, 028302.
Koster, D. A., Croquette, V., Dekker, C., Shuman, S. & Dekker, N. H. (2005). Friction and torque govern the relaxation of DNA supercoils by eukaryotic topoisomerase IB. 434, 671-4.
Kuscu, C., Arslan, S., Singh, R., Thorpe, J. & Adli, M. (2014). Genome-wide analysis reveals characteristics of off-target sites bound by the Cas9 endonuclease. 32, 677-83.
Lawrence, C. M. & White, M. F. (2011). Recognition of archaeal CRISPR RNA: No P in the alindromic repeat? 19, 142-4.
Leenay, R. T., Maksimchuk, K. R., Slotkowski, R. A., Agrawal, R. N., Gomaa, A. A., Briner, A. E., et al. (2016). Identifying and Visualizing Functional PAM Diversity across CRISPR-Cas Systems. 62, 137-47.
Lipfert, J., Kerssemakers, J. W., Jager, T. & Dekker, N. H. (2010). Magnetic torque tweezers: measuring torsional stiffness in DNA and RecA-DNA filaments. 7, 977-80.
Luzzietti, N., Brutzer, H., Klaue, D., Schwarz, F. W., Staroske, W., Clausing, S., et al. (2011). Efficient preparation of internally modified single-molecule constructs using nicking enzymes. 39, e15.
Maffeo, C., Schopflin, R., Brutzer, H., Stehr, R., Aksimentiev, A., Wedemann, G., et al. (2010). DNA-DNA interactions in tight supercoils are described by a small effective charge density. 105, 158101.
Makarova, K. S., Haft, D. H., Barrangou, R., Brouns, S. J., Charpentier, E., Horvath, P., et al. (2011). Evolution and classification of the CRISPR-Cas systems. 9, 467-77.
Makarova, K. S., Wolf, Y. I., Alkhnbashi, O. S., Costa, F., Shah, S. A., Saunders, S. J., et al. (2015). An updated evolutionary classification of CRISPR-Cas systems. 13, 722-36.
Mali, P., Yang, L., Esvelt, K. M., Aach, J., Guell, M., DiCarlo, J. E., et al. (2013). RNA-guided human genome engineering via Cas9. 339, 823-6.
Mojica, F. J., Diez-Villasenor, C., Garcia-Martinez, J. & Soria, E. (2005). Intervening sequences of regularly spaced prokaryotic repeats derive from foreign genetic elements. 60, 174-82.

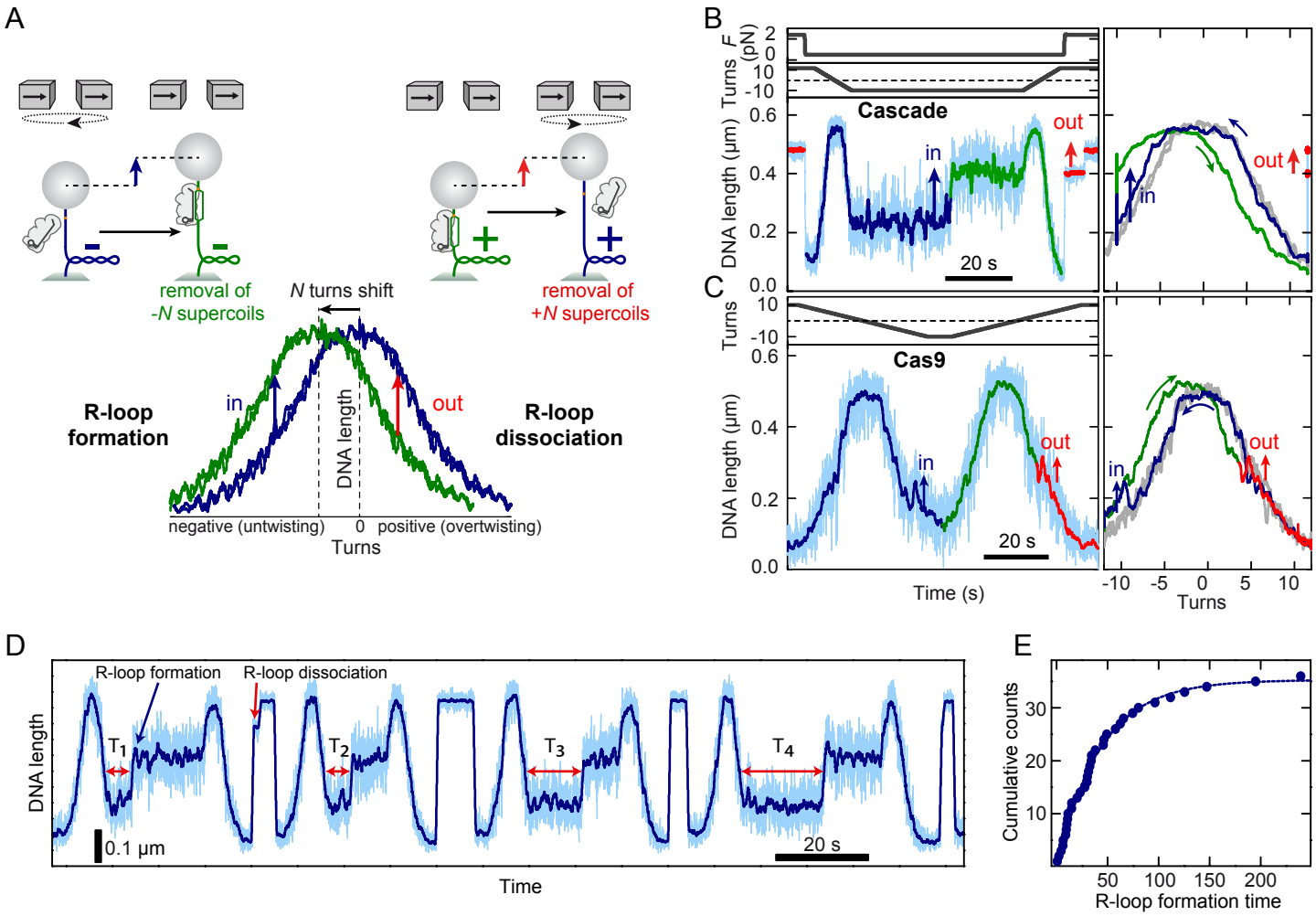
Mojica, F. J., Diez-Villasenor, C., Garcia-Martinez, J. & Almendros, C. (2009). Short motif sequences determine the targets of the prokaryotic CRISPR defence system. 155, 733-40.
Mosconi, F., Allemand, J. F., Bensimon, D. & Croquette, V. (2009). Measurement of the torque on a single stretched and twisted DNA using magnetic tweezers. 102, 078301.
Oberstrass, F. C., Fernandes, L. E. & Bryant, Z. (2012). Torque measurements reveal sequence-specific cooperative transitions in supercoiled DNA. 109, 6106-11.
Otto, O., Czerwinski, F., Gornall, J. L., Stober, G., Oddershede, L. B., Seidel, R., et al. (2010). Real-time particle tracking at 10,000 fps using optical fiber illumination. 18, 22722-33.
Pattanayak, V., Lin, S., Guilinger, J. P., Ma, E., Doudna, J. A. & Liu, D. R. (2013). High-throughput profiling of off-target DNA cleavage reveals RNA-programmed Cas9 nuclease specificity. 31, 839-43.
Pourcel, C., Salvignol, G. & Vergnaud, G. (2005). CRISPR elements in Yersinia pestis acquire new repeats by preferential uptake of bacteriophage DNA, and provide additional tools for evolutionary studies. 151, 653-63.
Ramanathan, S. P., van Aelst, K., Sears, A., Peakman, L. J., Diffin, F. M., Szczelkun, M. D., et al. (2009). Type III restriction enzymes communicate in 1D without looping between their target sites. 106, 1748-53.
Redding, S., Sternberg, S. H., Marshall, M., Gibb, B., Bhat, P., Guegler, C. K., et al. (2015). Surveillance and Processing of Foreign DNA by the Escherichia coli CRISPR-Cas System. 163, 854-65.
Revyakin, A., Ebright, R. H. & Strick, T. R. (2005). Single-molecule DNA nanomanipulation: improved resolution through use of shorter DNA fragments. 2, 127-38.
Revyakin, A., Liu, C., Ebright, R. H. & Strick, T. R. (2006). Abortive initiation and productive initiation by RNA polymerase involve DNA scrunching. 314, 1139-43.
Rouillon, C., Zhou, M., Zhang, J., Politis, A., Beilsten-Edmands, V., Cannone, G., et al. (2013). Structure of the CRISPR interference complex CSM reveals key similarities with cascade. 52, 124-34.
Rutkauskas, M., Sinkunas, T., Songailiene, I., Tikhomirova, M. S., Siksnys, V. & Seidel, R. (2015). Directional R-Loop Formation by the CRISPR-Cas Surveillance Complex Cascade Provides Efficient Off-Target Site Rejection.
Samai, P., Pyenson, N., Jiang, W., Goldberg, G. W., Hatoum-Aslan, A. & Marraffini, L. A. (2015). Co-transcriptional DNA and RNA Cleavage during Type III CRISPR-Cas Immunity. 161, 1164-74.
Sander, J. D. & Joung, J. K. (2014). CRISPR-Cas systems for editing, regulating and targeting genomes. 32, 347-55.
Schopflin, R., Brutzer, H., Muller, O., Seidel, R. & Wedemann, G. (2012). Probing the elasticity of DNA on short length scales by modeling supercoiling under tension. 103, 323-30.
Semenova, E., Jore, M. M., Datsenko, K. A., Semenova, A., Westra, E. R., Wanner, B., et al. (2011). Interference by clustered regularly interspaced short palindromic repeat (CRISPR) RNA is governed by a seed sequence. 108, 10098-103.
Sinkunas, T., Gasiunas, G., Fremaux, C., Barrangou, R., Horvath, P. & Siksnys, V. (2011). Cas3 is a single-stranded DNA nuclease and ATP-dependent helicase in the CRISPR/Cas immune system. 30, 1335-42.
Sinkunas, T., Gasiunas, G., Waghmare, S. P., Dickman, M. J., Barrangou, R., Horvath, P., et al. (2013). In vitro reconstitution of Cascade-mediated CRISPR immunity in Streptococcus thermophilus. 32, 385-94.
Sternberg, S. H., Redding, S., Jinek, M., Greene, E. C. & Doudna, J. A. (2014). DNA interrogation by the CRISPR RNA-guided endonuclease Cas9. 507, 62-7.
Strick, T. R., Croquette, V. & Bensimon, D. (2000). Single-molecule analysis of DNA uncoiling by a type II topoisomerase. 404, 901-4.
Strick, T. R., Allemand, J. F., Bensimon, D. & Croquette, V. (1998). Behavior of supercoiled DNA. 74, 2016-28.
Strick, T. R., Allemand, J. F., Bensimon, D., Bensimon, A. & Croquette, V. (1996). The elasticity of a single supercoiled DNA molecule. 271, 1835-7.

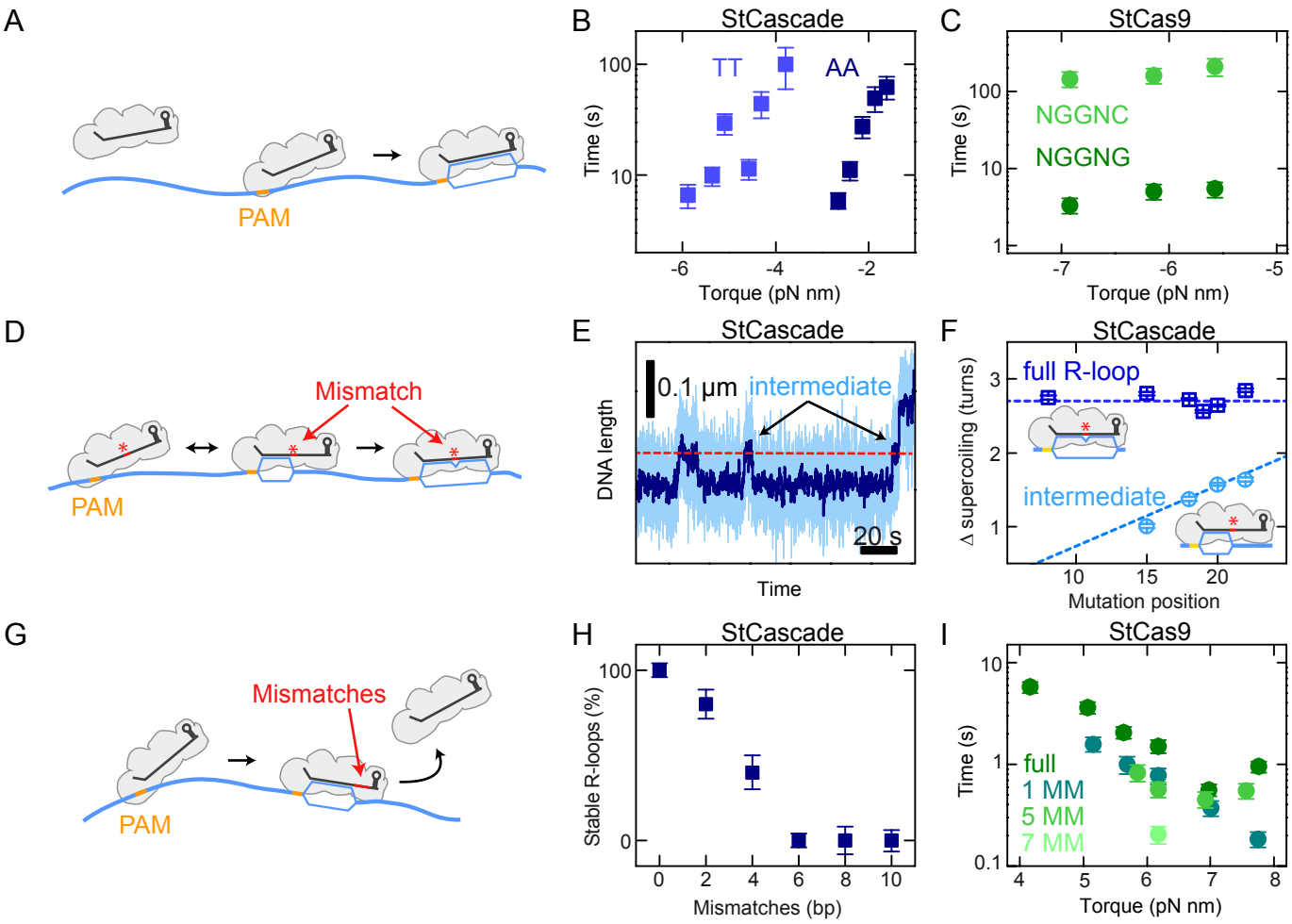
Szczelkun, M. D., Tikhomirova, M. S., Sinkunas, T., Gasiunas, G., Karvelis, T., Pschera, P., et al. (2014). Direct observation of R-loop formation by single RNA-guided Cas9 and Cascade effector complexes. 111, 9798-803.
Tamulaitis, G., Kazlauskiene, M., Manakova, E., Venclovas, C., Nwokeoji, A. O., Dickman, M. J., et al. (2014). Programmable RNA shredding by the type III-A CRISPR-Cas system of Streptococcus thermophilus. 56, 506-17.
van Aelst, K., Toth, J., Ramanathan, S. P., Schwarz, F. W., Seidel, R. & Szczelkun, M. D. (2010). Type III restriction enzymes cleave DNA by long-range interaction between sites in both head-to-head and tail-to-tail inverted repeat. 107, 9123-8.
van der Oost, J., Westra, E. R., Jackson, R. N. & Wiedenheft, B. (2014). Unravelling the structural and mechanistic basis of CRISPR-Cas systems. 12, 479-92.
Venable, R. M., Hatcher, E., Guvench, O., Mackerell, A. D., Jr. & Pastor, R. W. (2010). Comparing simulated and experimental translation and rotation constants: range of validity for viscosity scaling. 114, 12501-7.
Vlijm, R., Lee, M., Lipfert, J., Lusser, A., Dekker, C. & Dekker, N. H. (2015). Nucleosome assembly dynamics involve spontaneous fluctuations in the handedness of tetrasomes. 10, 216-25.
Wang, X., Wang, Y., Wu, X., Wang, J., Qiu, Z., Chang, T., et al. (2015). Unbiased detection of off-target cleavage by CRISPR-Cas9 and TALENs using integrase-defective lentiviral vectors. 33, 175-8.
Westra, E. R., Semenova, E., Datsenko, K. A., Jackson, R. N., Wiedenheft, B., Severinov, K., et al. (2013). Type I-E CRISPR-cas systems discriminate target from non-target DNA through base pairing-independent PAM recognition. 9, e1003742.
Westra, E. R., van Erp, P. B., Kunne, T., Wong, S. P., Staals, R. H., Seegers, C. L., et al. (2012). CRISPR immunity relies on the consecutive binding and degradation of negatively supercoiled invader DNA by Cascade and Cas3. 46, 595-605.
Wiedenheft, B., Lander, G. C., Zhou, K., Jore, M. M., Brouns, S. J., van der Oost, J., et al. (2011). Structures of the RNA-guided surveillance complex from a bacterial immune system. 477, 486-9.
Wu, X., Scott, D. A., Kriz, A. J., Chiu, A. C., Hsu, P. D., Dadon, D. B., et al. (2014). Genome-wide binding of the CRISPR endonuclease Cas9 in mammalian cells. 32, 670-6.
Yosef, I., Goren, M. G. & Qimron, U. (2012). Proteins and DNA elements essential for the CRISPR adaptation process in Escherichia coli. 40, 5569-76.
Zhang, J., Rouillon, C., Kerou, M., Reeks, J., Brugger, K., Graham, S., et al. (2012). Structure and mechanism of the CMR complex for CRISPR-mediated antiviral immunity. 45, 303-13.

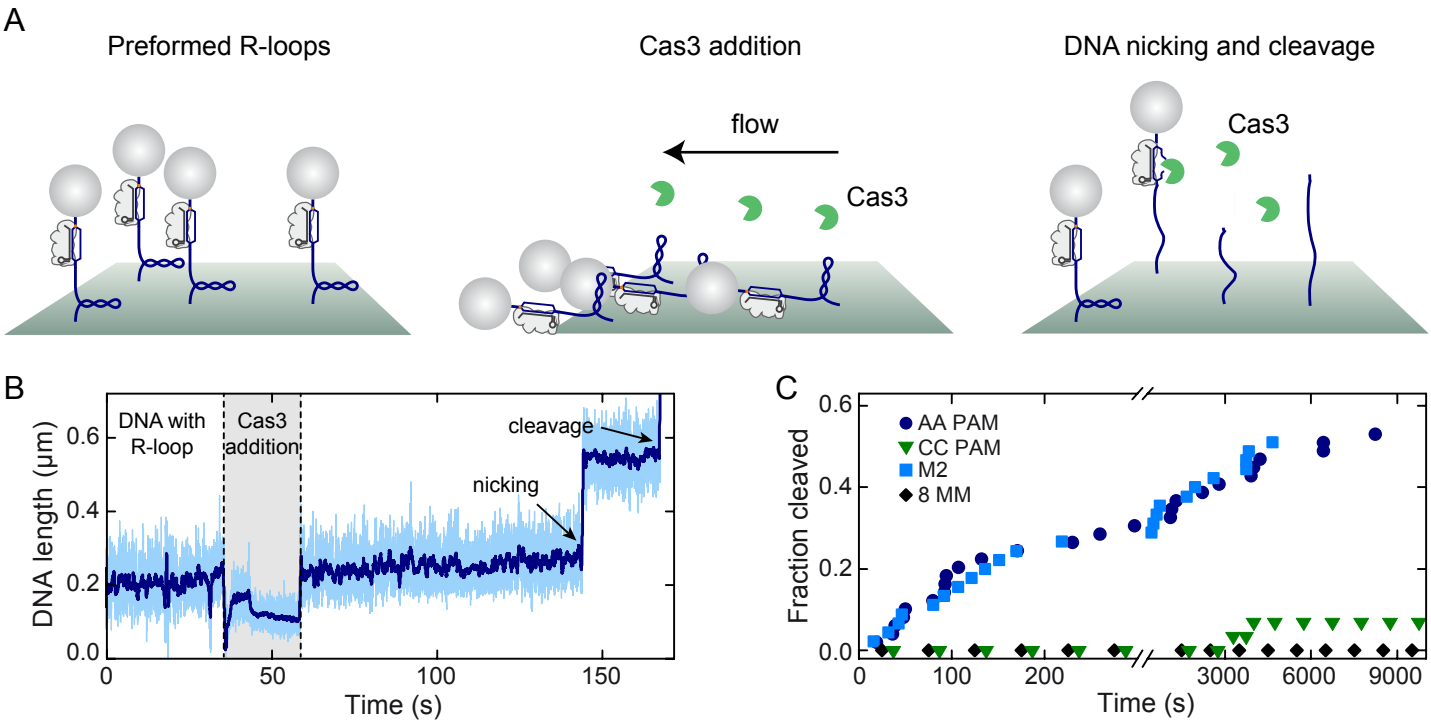

A
Adaptation
Cas genes
CRISPR locus
Transcription and
sRNAs maturation
Effector complex
ribonuclease
Interference
nuclease
B
Cas3
Cas7
Cas1 Cas2
Cas9
Cas1 Cas2
Cas10
Cas7
Cas1 Cas2
Type I
Type II
Type III

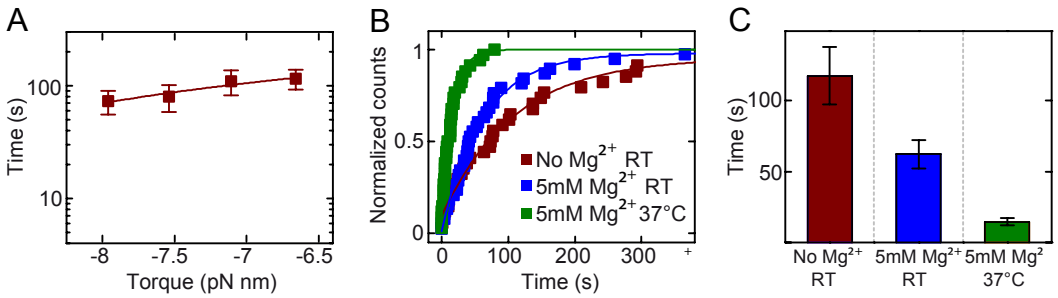

A
LED
Collimation
and focusing
lenses
Moveable permanent
magnets
Magnetic
field gradient
Flow cell
Objective
z
x
y
Tube lens
Camera
B
glass slide
parafilm
glass slide
C
Magnetic
bead
Biotin-streptavidin
interaction
Reference
bead
Digoxigenin-
anti-digoxigenin
interaction
Glass
slide
D
Diffraction ring patterns
-2 µm
-1 µm
Focus
+1 µm
+2 µm
Underfocus
Overfocus
E
x and y tracking
x
y
Intensity
F
z tracking
Increasing overfocus

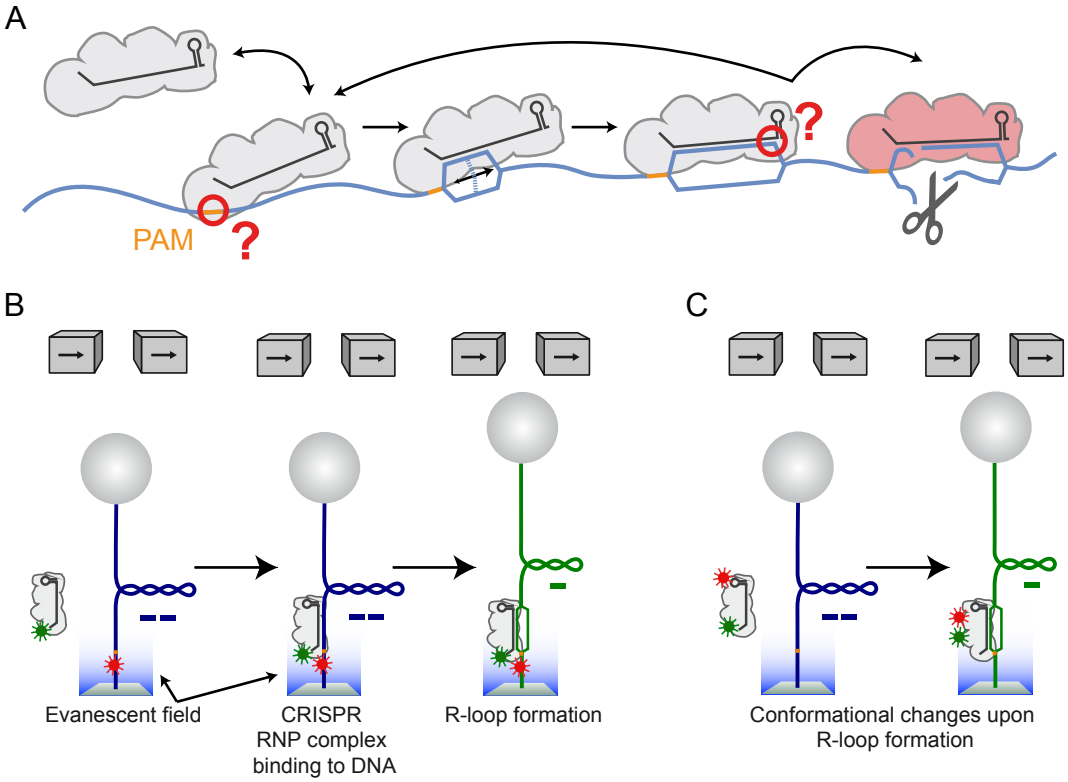

A
-
+
DNA length
negative (untwisting)
0
positive (overtwisting)
Turns
B
DNA length (μm)
0.6
0.4
0.2
Torque (pN nm)
30
20
10
5
10
15
Turns


A
removal of
-N supercoils
N turns shift
removal of
+N supercoils
R-loop
formation
in
DNA length
out
R-loop
dissociation
negative (untwisting)
0
positive (overtwisting)
Turns
B
F (pN)
Turns
DNA length (μm)
Cascade
in
out
20 s
C
Cas9
Time (s)
Turns
D
R-loop formation
R-loop dissociation
DNA length
T1
T2
T3
T4
0.1 μm
20 s
Time
E
Cumulative counts
R-loop formation time


A
PAM
B
StCascade
TT
AA
Time (s)
Torque (pN nm)
C
StCas9
NGGNC
NGGNG
Time (s)
Torque (pN nm)
D
Mismatch
PAM
E
StCascade
0.1 µm
intermediate
20 s
DNA length
Time
F
StCascade
full R-loop
intermediate
Δ supercoiling (turns)
Mutation position
G
Mismatches
PAM
H
StCascade
Stable R-loops (%)
Mismatches (bp)
I
StCas9
full
1 MM
5 MM
7 MM
Time (s)
Torque (pN nm)


A
Preformed R-loops
Cas3 addition
DNA nicking and cleavage
flow
Cas3
Cas3
B
DNA with
R-loop
Cas3
addition
cleavage
nicking
DNA length (μm)
0.0
0.2
0.4
0.6
0
50
100
150
Time (s)
C
AA PAM
CC PAM
M2
8 MM
Fraction cleaved
0.0
0.2
0.4
0.6
0
100
200
3000
6000
9000
Time (s)


A
Time (s)
100
10
-8
-7.5
-7
-6.5
Torque (pN nm)
B
Normalized counts
1
0.5
0
0
100
200
300
Time (s)
No Mg2+ RT
5mM Mg2+ RT
5mM Mg2+ 37°C
C
Time (s)
100
50
No Mg2+ RT
5mM Mg2+ RT
5mM Mg2 37°C


A
PAM
?
?
B
C
Evanescent field
CRISPR
RNP complex
binding to DNA
R-loop formation
Conformational changes upon
R-loop formation